



(19) Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) EP 1 085 756 A2

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
21.03.2001 Bulletin 2001/12

(51) Int. Cl.⁷: H04N 7/16

(21) Application number: 00120303.3

(22) Date of filing: 15.09.2000

(84) Designated Contracting States:

AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE

Designated Extension States:

AL LT LV MK RO SI

(30) Priority: 16.09.1999 US 154389 P
15.03.2000 US 525850

(71) Applicant:

Sharp Kabushiki Kaisha
Osaka-shi Osaka (JP)

(72) Inventors:

- van Beek, Peter
Rochester, NY 14620 (US)
- Sezan, Ibrahim
Vancouver, WA 98683 (US)
- Tomioka, Yoshiaki
Funabashi-shi, Chiba 274-0801 (JP)

8/91

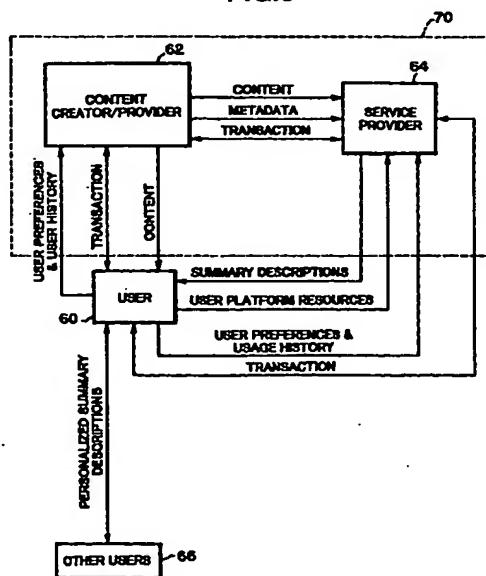
(74) Representative:

MÜLLER & HOFFMANN Patentanwälte
Innere Wiener Strasse 17
81667 München (DE)

(54) **Description framework for audiovisual content**

(57) A system and method for providing a description framework for an audiovisual presentation system. The system includes an interface that allows the user to consume several different representations of audiovisual content. The system also includes a descriptive structure that identifies and locates the summary selected by the user. The user is presented with a multi-view menu of the available types of summaries and selects a summary type, and the system provides summaries of that type to the user. The summary description service is provided to the user (60) based upon user preferences and history. When audiovisual material is then transmitted to the user (60), the description service provides the user (60) with the summary description that allows the user (60) to make and consume summaries of the material.

FIG.6



Description**BACKGROUND OF THE INVENTION**

5 [0001] This application is a continuation of US Provisional Application No. 60/154,389, filed 9/16/99 and claims priority thereto.

1. Field of the Invention

10 [0002] This invention relates to descriptions of audio-visual material. More particularly, this invention relates to a method for specifying descriptions that allow users to navigate amongst different audiovisual material, and browse and experience the content of a particular audiovisual program, quickly and effectively.

2. Background of the Invention

15 [0003] Digital audiovisual material is becoming increasingly available to users through digital TV broadcast, digital video cameras, DVD, and PC-based access to multimedia on the Internet. In addition, persistent large-volume storage and storage that allows non-linear access to audiovisual content, such as hard disk storage in powerful PC platforms and personal video recorders (PVR), is becoming available in consumer devices. Consequently, there is a need for
20 rapid navigation and browsing capabilities to enable users to efficiently discover and consume the contents of audiovisual material or programs.

[0004] Users would also benefit from having non-linear access to different views of a particular program, a feature not currently available. The views could be adaptive to user's personal preferences, interests or usage conditions, such as the amount of time the user wants to spend in consuming the content, or the resources available to the user's terminal. Such adaptability would enhance the entertainment and educational value of audiovisual information.

[0005] This proliferation of audio-visual material available to users has the potential to overwhelm the viewer and lead to frustration at the inability to browse and view content in an efficient manner. Viewing summaries of the content allows the viewer/user to skip irrelevant content and locate the desired content quickly and easily. Further, multiple different summaries, if available, could provide the user with alternative views of a particular program that the user could
30 choose from depending on personal preferences or usage conditions.

[0006] This capability is appearing more frequently in newer technologies, such as the digital video disk (DVD). DVD movies provide 'scene selections' or 'chapter selections' that have a visual array of thumbnails and textual titles associated with each scene. This allows the user to click on the thumbnail of the desired scene, jump to that scene and begin playback. Playback typically continues until the end of the movie, unless the viewer makes another selection.

[0007] However, this technology remains limited, providing only the capability to index for the purpose of jumping to an arbitrary position and continuing playback from that position. Additionally, these are only currently available for movies and cannot be provided for other types of audio-visual content, such as home movies, or recordings of real-time broadcast of television. This capability can be seen as a visual index, a simple form of a summary description.

[0008] A system in which such summaries and descriptions can be used is discussed in co-pending US Patent
40 Application No. 09/299,811, filed 4/26/99, and owned by the assignee of this application and incorporated by reference herein. The system discussed functions in a typical audiovisual system including several devices such as a television, cable or satellite reception, a sound system, etc. The term system refers to both individual devices and systems of several of these devices.

[0009] However, the reference does not provide certain aspects of implementation of such a system, including
45 models for usage and provision of content and services.

SUMMARY OF THE INVENTION

[0010] One aspect of the invention is a system that provides a descriptive framework about programs presented by
50 an audiovisual system. The framework includes an interface allowing a user to view representations of audiovisual material and a descriptive structure that identifies and locates each of the representations of audiovisual material and data associated with the representation. Examples of such representations could be a multimedia title description and summary descriptions.

[0011] Another aspect of the invention is a method for providing alternative summaries to the user having the steps
55 of presenting a multi-view menu of the available types of summaries to the user. The summaries can be hierarchical or non-hierarchical. A user selection of a summary type is received and the selected summaries are provided.

[0012] Yet another aspect of the invention is a method of providing summary description services of audiovisual content to a user. Information is received from the user including specifications of platform resources at the user end

and user preferences. Usage history of the user is tracked and used in conjunction with the specifications and preferences to transmit audiovisual material to the user with associated summary descriptions. The summary descriptions can be provided in such a way that the user can send the summary descriptions to other users.

5 BRIEF DESCRIPTION OF THE DRAWINGS

[0013] For a more complete understanding of the present invention and for further advantages thereof, reference is now made to the following Detailed Description taken in conjunction with the accompanying Drawings in which:

10 Figure 1 shows a block diagram representation of an audiovisual presentation system in accordance with the invention.

Figure 2 shows a method for selection of an audiovisual program within a description framework in accordance with the invention.

15 Figure 3 shows a block diagram representation of a summary description scheme in accordance with the invention.

Figure 4 shows a block diagram representation of alternative summaries available within a summary description scheme in accordance with the invention.

20 Figure 5 shows a block diagram representation of alternative summaries available within a hierarchical summary description scheme in accordance with the invention.

25 Figure 6 shows a flow chart for one embodiment of provision of audiovisual services in accordance with the invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0014] As mentioned previously, an overall system for creating and managing description schemes for an audiovisual presentation system is disclosed in co-pending US Patent Application No. 09/299,811, filed 4/26/99.

30 [0015] In this system, the video, image and/or audio information, which will be referred to as presentation information is made available to a user and/or a system. The information is presented to the user from the system such as a television or radio. The user or user's agent interacts with the system to receive the information in a desirable manner and to define preferences as to what type of information is obtained. The term user will refer to the end recipient of the information, which could be a person, a machine or a software program running on a machine, as examples.

35 [0016] To define these interactions, a set of description schemes containing data describing each component is defined, with reference to an overall audiovisual presentation system 10 as shown in Figure 1. The user preferences 12 are used in several different areas to maximize both the user's enjoyment and the system utility to the user. The preferences describing the topics and subject matter of interest to the user is used in both searching for and filtering the audiovisual programs 14. These two sets of data, the user preferences and program descriptions 14, are correlated in the filtering and search, engine 16 to identify the preferred programs.

40 [0017] The programs identified by the filtering and search engine 16 is then sent to a browsing module 18, along with the user's browsing preferences. Another output from the filtering and search engine 16 are preferred programs that the user has designated for storage. These are stored in storage module 20. The programs selected by the user 45 with the browsing module are then sent to the display 22. The user may utilize multimedia title descriptions of preferred programs to navigate among the programs that the user wants to consume. Once a program is selected, summary description of that particular program is correlated with user's browsing preferences to offer the user preferred alternative summaries.

45 [0018] The display 22 receives the programs and displays them in accordance with the user's device preferences as to the operation of the display. User's device preferences may include, for example, device settings such as volume setting that may vary with the genre of the program that is being watched. The display and user's interaction with the display, such as stopping a program before its end and watching certain types of programs with certain device settings, also provides information in a manner analogous to a feedback loop to update and log the usage history 24. The usage history 24 can be mapped against the preferences by mapping module 26. This information is then used in conjunction 55 with user inputs by the user preference module 12.

[0019] These documented user preferences can be useful in several contexts, not just an audiovisual presentation system. The user preferences and usage history conform to a specified format similar to that of the description of the audiovisual program information and can therefore be accumulated in the system as usage history information for fur-

ther use in selecting the contents desired by the user. Furthermore, the usage history information can be transmitted to the provider of the audiovisual programs 14 to receive a selected audiovisual program or directly receive audiovisual program summaries selected by the user. In the latter case, user preferences are correlated with summary descriptions at the provider side to select and directly deliver summarized audiovisual programs to the user. The preferences and summary descriptions and so on could also be transferred to a 'smart card' 28 or similar, portable storage means and ultimately transferred to another system by the user. However, the details of this type of transfer are beyond the scope of the current invention and are only mentioned for completeness.

[0020] In this context, there are several description schemes, which were the subject matter of the previously mentioned co-pending patent application. However, that discussion was at the overall system level, and no framework for the individual descriptions schemes at the user, program or device level were considered. This invention is intended to provide, as a technique of realizing efficient navigation and browsing of audiovisual programs using their summaries, a title description scheme capable of including multimedia information and a summary description scheme for describing hierarchical summaries of an audiovisual program and, furthermore, to construct a system and a service model utilizing description data based on the above description schemes.

[0021] As shown in Figure 1, the audiovisual programs 14 include descriptions of the programs in a description framework. The description framework can have several different types of descriptive structures. Of particular interest here are the multimedia title description and the summary description. The framework can contain either one of these types, both of these and either one or both in combination with other types of program descriptions including metadata. For example, metadata on the creation of the program (director, actors, language, etc.), and genre of the program can be provided.

[0022] In operation, the user manipulates the descriptive structures to select audiovisual programs presented by the presentation system of Figure 1. This view and select process can occur in several ways, as shown in Figure 2. For ease of discussion and understanding, one could view the description framework like an electronic library. The user could browse and search the programs by their titles, analogous to the multimedia title description, or by a more robust summary, analogous to the summary descriptions. The descriptive structures such as the multimedia title description and the summary descriptions can be in one of several forms, including text, audio clips, video clips, still images, etc.

[0023] Summary descriptions enable rapid navigation and browsing in this system. In particular, summary descriptions enable key-frame summaries, event based video summaries that group video segments containing certain events, and video highlight summaries of particular duration. These summary descriptions's schemes contain references to the audiovisual media and its segments, frames, and audio tracks that can be efficiently utilized by a presentation engine in rendering different summaries and views of a program. Hence, when an audiovisual program has multiple versions of its summary description, the system can subsequently generate respective views by means of the presentation engine using each of the summary description versions. Consequently, the system provides an efficient means for using multiple views of a program without the need for pre-storing its multiple versions in a separate storage area, thus realizing saving in an area for storing data at the system side.

[0024] The term summary description as used here refers to summaries that conform to a set of rules for such summaries. The syntax, semantics, and rules of for these summary descriptions are contained in summary description schemes. Summary description schemes specify which descriptors and attributes can be used in the description, their allowed range of values and the rules for their combinations. The use of common set of description schemes and descriptors would enable interoperability between different devices (i.e., devices for content providers, devices for content creators, devices for service providers and devices for users) that handle audio-visual content. These different devices would all be able to interpret summaries that use the same description schemes and descriptors. Ideally, the scheme would allow the different devices flexibility in how it presents the contents of the summaries to the viewer.

[0025] A particular audiovisual program today is often created rich in media. In particular, it may have a still image, graphic, short video clip, an audio jingle, or a pictorial logo associated with it, which concisely represent its content. Such media can be used along with the usual textual title of a program. For example, a music program may have a pictorial title in addition to its textual title; a TV program may have a logo or an audio jingle. Fig. 3 illustrates an example of a description scheme for integrating the data such as a text, an audio clip, a video clip and a still image into information associated with a title. In the descriptive structure of Fig. 3, a title is described inside (Title Text).../(Title Text) like a title in conventional text data, while information for locating multimedia data such as an audio clip, a video clip and a still image is described inside (Title Image).../(Title Image). This creates a description scheme enabling collection of the conventional text data and the information for locating multimedia data. The use of the description scheme enables the system to extract, for example, image data from a storage area for storing multimedia such as image, develop and add the data. Namely, the system not only presents text data but also prepares a multimedia title easily. The multimedia title descriptions can facilitate an audiovisual, informative, effective, and entertaining navigation between different audiovisual programs.

8, 9,

[0026] In this title description scheme, the multimedia data is represented by information for locating the multimedia data. This enables the system to prepare multimedia data information not only by directly referring to the multimedia

data stored in the system but also by specifying a frame number of an AV program stored in the system, specifying the beginning time and the ending time of a video clip and an audio clip or referring to multimedia data being at the side of provider outside the system.

5 [0027] Consequently, the system that does not contain original multimedia information can prepare a multimedia title description by referring the location of the multimedia data in the description data. This eliminates the need for storing the multimedia information at the system side, achieving the saving in data storage area thereof. If the data stored outside the system is necessarily used multiple times, the system can read first and then use the data as the internal data to assure rapid presentation of the data.

10 [0028] Multimedia title description may be integrated into the summary description, or may be a separate description for a particular program, without impacting the functionality of multiple media titles and summaries. Once a user chooses a program of interest as a result of navigation through multimedia titles, the user may utilize the summary description for that program to quickly discover the audiovisual content of the program, browse the program, or non-linearly navigate within the program. The relationship between these descriptive structures will be discussed with reference to Figure 2.

15 [0029] As shown in Figure 2, the user can start by experiencing the multimedia title descriptions at step 30. The user can then select a title and hence an audiovisual program and go to the next level of description of that audiovisual program, which would be the summary description in step 32. The user then makes a selection in step 34 and browses and experiences the program.

20 [0030] Alternatively, the user could skip the summary description in step 32 and make a selection based only on the multimedia title description. Another possibility allows the user to skip viewing the multimedia title descriptors and instead starting the selection process at the summary description level. The starting point for the user may be determined by the amount of time available, any previous knowledge of the programs, and the desired amount of detail. Similarly, it is also possible to adaptively read first only description data such as multimedia title data and summary data into the user system at Step S30 and obtain necessary audiovisual program data through Step S32.

25 [0031] An example of a multimedia title description is shown below.

```

<Title>
  <TitleText>
    Afternoon news
  </TitleText>
  <TitleImage>
    <MediaURL>file://thumbnails/news001.jpg</MediaURL>
  </TitleImage>
</Title>

```

30
35

40 [0032] In order to understand the higher amount of detail used in the summary description, it is helpful to discuss it more thoroughly. As shown in Figure 3, audiovisual summaries are extracted from the audiovisual media 48 by the extraction module 46. Descriptions of these summaries 44 are then authored according to the summary description scheme which specifies the elements, descriptors, attributes and other descriptions that can be included in the description, the ranges of values that they can attain, and their allowable combinations of the elements, description and attributes. The summary description scheme includes a data description scheme for preparing a multimedia title and a description scheme for presenting a summary description of an audiovisual program. The multimedia title description and the summary description of an audiovisual program are prepared according to the above scheme. The displaying device 42, whether a user terminal or audiovisual device, receives the summary description and the audiovisual content. The device 42 includes a parser 50 that interprets and validates the audiovisual summary description 44 in accordance with the description scheme and presents the summaries to the user with the interface 52.

45 [0033] The summary description scheme of Figure 3 is shown in more detail in Figure 4. The content of Fig. 4 is shown in more detail in Fig. 5. To satisfy the various kinds of requirements from users, content providers and service providers, the present invention provides description schemes capable of describing a variety of summary descriptions as shown later. By using the description schemes, the provider side and the user side may have a common framework. Hence, the user can select a desired summary description by utilizing the summaries provided by the provider. Exemplified alternatives selectable by the user can be structured by hierarchically arranging a variety of descriptions such as "want to see a 10-second highlight scene", "want to see only a slam dunk shot" and the like. Furthermore, the provider can provide the user with a plurality of summary descriptions to meet the user preferences using the user usage history information received from the user. These summary descriptions are structured to contain references to the audi-

ovisual media and its segments, frames, and audio tracks that can be efficiently utilized by a presentation engine in rendering different summaries and views of the program. The example of Fig. 4 illustrates in detail a summary description scheme used for representing a variety of summaries as above described. The attribute 'summary type' 41 defines the type of summary description scheme 40. The use of this attribute enables the user to select either a hierarchical summary description scheme 54 or a sequential summary description scheme 56. The attributes of highlight or multiresolution are both constructed with a hierarchical description scheme, shown in more detail in Figure 5. The hierarchical summary description scheme 54 will be described later.

[0034] Figure 5 is a block diagram illustrating the hierarchical summary description schemes selectable for preparing summary descriptions. The scheme is as follows: The hierarchical summary description 54 is used to specify and group summaries of an audiovisual program, which may be structured hierarchically. It contains description data of a technique for constructing a hierarchy of an attribute 'summary type'. The hierarchical summary description 54 has plural hierarchical summary level descriptions. The hierarchical summary level descriptions are labeled and organized at different levels as shown below. Each level describes a summary of the audiovisual program by information at a specific level. The hierarchical summary level description is structured in such a way that it may have a further hierarchical summary level description to define a further deep level summary. In general, levels closer to the root of the hierarchy provide coarse summaries and levels further away from the root provide more detailed summaries.

```

20 <highlightSummary summaryType="highlight">
    <highlightSegmentLevel name="Level1_Coarse">
        <highlightSegmentLevel name="Level2_Middle">
            </highlightSegmentLevel>
        <highlightSegmentLevel name="Level2_Middle">
            <highlightSegmentLevel name="Level3_Fine">
                </highlightSegmentLevel>
            </highlightSegmentLevel>
        </highlightSegmentLevel>
    </highlightSegmentLevel>
</highlightSummary>
30

```

[0035] The hierarchical summary description thus structured enables the detailed summary description to include the coarse summary description, eliminating the duplication of the same data for representing the summaries. To view an audiovisual program using the summary description, the user can operate the presentation engine using a desired summary level description and a higher-level summary description.

As shown in Fig. 5, the hierarchical summary level description contains references to audiovisual media and its segments, frames and audio tracks. The reference may be made to segments and frames inside and outside the system. Hence, the system by itself can obtain audiovisual media data from external data stored at external providers and/or 40 inside the system in accordance with the selected summary and can provide the user with multiple views by using the presentation engine. When preparing a summary of a trailer for a serial film program, the description may contain a location of storing the proceeding audiovisual program and its highlight scene and time duration in addition to the preceding audiovisual program and its highlight scene and time duration. This offers the advantage of saving data storing areas at both the provider side and the user.

45 The above hierarchical summary description has an attribute 'Hierarchy Type', which specifies the type of interrelation between different levels of the summary. The attribute value can specify whether the hierarchy type is dependent or independent. If the hierarchy type is "independent", the information in a hierarchical summary level can completely specify a particular summary, without reference to the information in its parent element. However, it has such a demerit that an amount of necessary data is increased. If hierarchy type is "dependent", the summary at a particular level cannot be prepared without knowledge of its parent element. In this case, a demerit of necessarily organizing data into a 50 hierarchical system is caused but a merit of reducing the data amount is obtained. The user who desires a summary may select either of the types in accordance with the system specification. There is a description scheme for integrating all the summaries having the above-described features.

[0036] Audiovisual summaries enable users to consume alternative views of a particular audiovisual program 55 where views can be chosen according to the amount of time available to the user, personal preferences and point of views, and amount of resources available to user's platform. These are achieved by summary descriptions using the above-described summary description schemes.

[0037] In a hierarchical summary context where summaries can be at different hierarchy levels, grouping enables

summaries to be at the same hierarchy level and made available to the user as alternatives at that same level. For example, a multiview highlight summary may have a 30-second and a 60-second level for time constrained viewing at different detail/length levels. On the other hand, a multiview event summary enables summaries based on different events and different points of view where these summaries do not necessarily have a hierarchical relationship among them and thus they are merely alternatives.

5 [0038] The names of the alternative summaries can be presented to the user in an interactive menu and the user selects the desired summary by using summary description data.

10

ALTERNATIVE SUMMARIES

THREE POINT SHOTS

SLAM DUNK

15

...

....

MY FAVORITE MOMENTS

20

[0039] An example description according to the above-described summary description schemes that will support the items in this menu is given below. Note that the above menu items correspond to the "Highlight Level Name" descriptors in the description. Indeed, the description may utilize a numerical, machine-readable code corresponding to the string that is presented to the user by the presentation engine. This is an implementation issue.

25

30

35

40

45

50

55

```

<Program>
  <MediaInformation>
    <MediaProfile>
      <MediaInstance><Locator>file://disk/sports/video1.mpg</Locator></MediaInstance>
    </MediaProfile>
  </MediaInformation>
  <MetaInformation>
    <Creation><Title>Blazers vs Pacers 2/10/99</Title></Creation>
    <Classification><Genre>Sports</Genre></Classification>
  </MetaInformation>
  <Summarization>
</HierarchicalSummary>
  <HierarchicalSummary summaryType="keyEvents" name="Multiview event summary">
    <HighlightLevel name="Three Point Shots">
      <HighlightSegment name="Three Point Shot #1">
        <VideoSegmentLocator><MediaTime>10410 10680</MediaTime></VideoSegmentLocator>
        <ImageLocator><MediaTime>10800</MediaTime></ImageLocator>
      </HighlightSegment>
      ...
      <!-- more video-segments -->
    </HighlightLevel>
    <HighlightLevel name="Slam Dunks">
      <HighlightSegment name="Slamdunk #1">
        <VideoSegmentLocator><MediaTime>13350 13560</MediaTime></VideoSegmentLocator>
        <ImageLocator><MediaTime>13500</MediaTime></ImageLocator>
      </HighlightSegment>
      ...
      <!-- more video-segments -->
    </HighlightLevel>
    <HighlightLevel name="My Favorite Moments">
      <HighlightSegment name="The Best 2-Point Shot from Blazers">
        <VideoSegmentLocator><MediaTime>10110 10210</MediaTime></VideoSegmentLocator>
        <ImageLocator><MediaTime>10180</MediaTime></ImageLocator>
      </HighlightSegment>
      <HighlightSegment name="Scotty Pippin's Best Basket">
        <VideoSegmentLocator><MediaTime>16110 17210</MediaTime></VideoSegmentLocator>
        <ImageLocator><MediaTime>16903</MediaTime></ImageLocator>
      </HighlightSegment>
      ...
      <!-- more video-segments -->
    </HighlightLevel>
  </HierarchicalSummary>
  <Summarization>
</Program>

```

[0040] Note that the menu above can facilitate bookmarking of a multitude of segments grouped under one theme, which in this case is an event-based theme. The grouping (or bookmarking) is at a "highlight level" and is at the same

hierarchical level. The last item in the menu ("My Favorite Moments") corresponds to a fragment of the description that might have been generated by the user utilizing an appropriate authoring tool in his/her system. In other words, the user may have necessary tools to augment a summary that is available from a service provider, or tools to create a summary description from scratch.

- 5 [0041] Segment-level bookmarking is also possible as seen in the last part (shaded) of the description shown above. The user in the above example has marked two segments as "The Best 2-Point Shot from Blazers" and "Scotty Pippin's Best Basker". These two bookmarks can be presented to the user in the form of either a separate menu of bookmarks, or a sub-menu of the menu item "My Favorite Moments".
- 10 [0042] Alternative summaries that are not necessarily hierarchically structured are also allowed, e.g., a summary containing clips of "goals" vs. a summary contain clips of "passing shots". Such grouping of summaries is necessary to allow different event views.
- [0043] A hierarchical summary refines a summary and is the root element of the hierarchical summary. A hierarchical summary may contain multiple hierarchical summary level elements as shown below. Each hierarchical summary level element specifies a (hierarchical) summary and groups a number of video segments. These summaries represent
- 15 alternative views of the video program.
- [0044] A hierarchical summary description has an attribute hierarchyType, which specifies the type of interrelation between different levels of the summary. The hierarchyType can be independent or dependent. If hierarchyType is independent, the information in a hierarchical summary level completely specifies a particular summary, without reference to the information in its parent element. If hierarchyType is dependent, information in a hierarchical summary level adds
- 20 to, or refines, the information in its parent element; i.e., the summary at a particular level can't be reconstructed without knowledge of the parent element.
- [0045] The following is an example of a hierarchical summary according to the above-described summary description schemes that contains a highlight summary. The definition of highlight level is given below. The highlight summary may, for example, contain interesting video clips, ordered in multi-level fashion. Since the hierarchyType of this Hierarchical summary is dependent, a highlight summary at level $n+1$ adds more video clips to the highlight summary at level n . Thus, each level accumulates more information to provide a longer and more extensive video summary.

(Hierarchical summary name="mySummary" hierarchyType="dependent")

- 30 (HighlightLevel).../(HighlightLevel)
- /Hierarchical summary)

- [0046] The hierarchical summary level description scheme is used to specify a summary at a particular level of detail. The hierarchical summary level description scheme is an abstract scheme from which two types of summary description schemes are derived, either a highlight level description scheme, or a multiresolution description scheme. Multiresolution description schemes are outside the scope of this disclosure and is only included for completeness. The hierarchical summary level may contain zero or more hierarchical summary level elements as its children.
- [0047] As mentioned above the highlight level description scheme is used to specify a summary by referring to a sequence of audio-visual segments and their key-frames. A highlight level refines a hierarchical summary level and contains a single summary or part of a summary. A highlight level contains a sequence of references to video segments and their representative key-frames. A locator specifies each video segment and another locator specifies each representative key-frame. A highlight level has a required attribute flame, and an attribute level, which specifies the level of this summary in the hierarchy. It also has an attribute duration, which specifies the total duration of the summary at the same level in the hierarchy.
- [0048] The following is an example of a simple highlight summary according to the above-described summary description schemes with a duration of 10 seconds. It consists of two video clips, the first from frame 0 to 120 and the second from frame 200 to 380. The key-frame for the first video clip is frame 60 and the key-frame for the second clip is frame 320. Note that a key-frame may be a frame that is visualized before the video segment itself is played; e.g., playback of the video segment is activated by the user clicking on the key-frame.

```

5   <highlightLevel name="10 second highlight" duration="10">
10    <highlightSegment>
15     <videoSegmentLocator>
20      <mediaTime>0 120</mediaTime>
25     </videoSegmentLocator>
30     <imageLocator>
35       <mediaTime>60</mediaTime>
40   </highlightSegment>
45   </highlightLevel>
50 </highlightSummary>

```

[0049] The following is an example of a set of two summaries according to the above-described summary description schemes, one being an alternative summary to the other. The first highlight summary is 4 seconds long and contains only a single video clip, while the second summary is 10 seconds long and contains three video clips. Note that both summaries share one video clip; there is a hierarchical structure evident in the underlying data, yet its representation is not hierarchical.

```

35   <highlightLevel name="4 second highlight" duration="4">
40    <highlightSegment>
45     <videoSegmentLocator>
50       <mediaTime>1000 1120</mediaTime>
55     </videoSegmentLocator>
60     <imageLocator>
65       <mediaTime>1060</mediaTime>
70     </imageLocator>
75   </highlightSegment>

```

```

</HighlightLevel>
<highlightLevel name="10 second highlight" duration="10">
    <highlightSegment>
        <VideoSegmentLocator>
            <MediaTime>200 290</MediaTime>
        </VideoSegmentLocator>
        <ImageLocator>
            <MediaTime>200</MediaTime>
        </ImageLocator>
    </highlightSegment>
    <highlightSegment>
        <VideoSegmentLocator>
            <MediaTime>1000 1120</MediaTime>
        </VideoSegmentLocator>
        <ImageLocator>
            <MediaTime>1050</MediaTime>
        </ImageLocator>
    </highlightSegment>
    <highlightSegment>
        <VideoSegmentLocator>
            <MediaTime>1200 1290</MediaTime>
        </VideoSegmentLocator>
        <ImageLocator>
            <MediaTime>1200</MediaTime>
        </ImageLocator>
    </highlightSegment>
    <highlightSegment>
        <VideoSegmentLocator>
            <MediaTime>1200 1290</MediaTime>
        </VideoSegmentLocator>
        <ImageLocator>
            <MediaTime>1200</MediaTime>
        </ImageLocator>
    </highlightSegment>
</HighlightLevel>
</Hierarchical summary>

```

[0050] The following is an example of the same set of two summaries according to the above-described summary
40 description schemes, one being an alternative summary to the other. However, they are now hierarchically represented in the description, such that the application is informed of the underlying hierarchy in the data. In this example, the hierarchy type is "independent", which means that the common video clip must be repeated on the finer level, because the information on the finer level must be interpreted independently.

```

45 <Hierarchical summary name="mySummary"
      summaryType="highlight" hierarchyType="independent">
    <highlightLevel name="4 second highlight" duration="4">
        <highlightSegment>

```

```

<VideoSegmentLocator>
  <MediaTime>1000 1120</MediaTime>
</VideoSegmentLocator>
<ImageLocator>
  <MediaTime>1060</MediaTime>
</ImageLocator>
</HighlightSegment>
<HighlightLevel name="10 second highlight" duration="10">
  <HighlightSegment>
    <VideoSegmentLocator>
      <MediaTime>200 290</MediaTime>
    </VideoSegmentLocator>
    <ImageLocator>
      <MediaTime>200</MediaTime>
    </ImageLocator>
  </HighlightSegment>
  <HighlightSegment>
    <VideoSegmentLocator>
      <MediaTime>1000 1120</MediaTime>
    </VideoSegmentLocator>
    <ImageLocator>
      <MediaTime>1060</MediaTime>
    </ImageLocator>
  </HighlightSegment>
  <HighlightSegment>
    <VideoSegmentLocator>
      <MediaTime>1200 1290</MediaTime>
    </VideoSegmentLocator>
    <ImageLocator>
      <MediaTime>1200</MediaTime>
    </ImageLocator>
  </HighlightSegment>
</HighlightLevel>
</Hierarchical summary>

```

[0051] The following is a more complex example of a hierarchical summary according to the above-described summary description schemes consisting of the same video clips, organized into two levels. At the highest level, the summary has a duration of 4 seconds and consists of only one video clip. At the second level, the summary has a duration of 10 seconds and consists of three video clips. Note that both summaries again share one video clip as in the previous example, but it is specified only once, by utilizing a dependent hierarchical representation (hierarchy Type is dependent).

```

<HierarchicalSummary name="mySummary"
    summaryType="highlight" hierarchyType="dependent">
    <HighlightLevel name="4 second highlight" duration="4">
        <HighlightLevel name="10 second highlight" duration="10">
            <HighlightSegment>
                <VideoSegmentLocator>
                    <MediaTime>200 290</MediaTime>
                </VideoSegmentLocator>
                <ImageLocator>
                    <MediaTime>200</MediaTime>
                </ImageLocator>
            </HighlightSegment>
        </HighlightLevel>
        <HighlightSegment>
            <VideoSegmentLocator>
                <MediaTime>1000 1120</MediaTime>
            </VideoSegmentLocator>
            <ImageLocator>
                <MediaTime>1060</MediaTime>
            </ImageLocator>
        </HighlightSegment>
        <HighlightLevel name="10 second highlight" duration="10">
            <HighlightSegment>
                <VideoSegmentLocator>
                    <MediaTime>1200 1290</MediaTime>
                </VideoSegmentLocator>
                <ImageLocator>
                    <MediaTime>1200</MediaTime>
                </ImageLocator>
            </HighlightSegment>
        </HighlightLevel>
    </HighlightLevel>
</HierarchicalSummary>

```

[0052] The following is an example of a different set of two summaries, one being an alternative summary to the other, ordered in a two-level hierarchy. In this case, the video clips on the finer level are sub-clips of the single clip on the coarse level. By utilizing the hierarchical representation, the application is informed there is some type of hierarchical relation between the two summaries. However, the hierarchy type is "independent", since the clips on the finer level do not literally include the clip on the coarse level (the information at the coarse level cannot be reused).

```

<Hierarchical summary name="mySummary"
    summaryType="highlight" hierarchyType="Independent">
    5
    <HighlightLevel name="10 second summary" duration="10">
        <HighlightSegment>
            <VideoSegmentLocator>
                <MediaTime>1000 1300</MediaTime>
            10
            </VideoSegmentLocator>
            <ImageLocator>
                <MediaTime>1060</MediaTime>
            </ImageLocator>
        15
        </HighlightSegment>
        <HighlightLevel name="10 second summary" duration="10">
            <HighlightSegment>
                <VideoSegmentLocator>
                    <MediaTime>1000 1090</MediaTime>
                20
                </VideoSegmentLocator>
                <ImageLocator>
                    <MediaTime>1030</MediaTime>
                </ImageLocator>
            25
            </HighlightSegment>
            <HighlightSegment>
                <VideoSegmentLocator>
                    <MediaTime>1090 1210</MediaTime>
                30
                </VideoSegmentLocator>
                <ImageLocator>
                    <MediaTime>1120</MediaTime>
                </ImageLocator>
            35
            </HighlightSegment>
            <HighlightSegment>
                <VideoSegmentLocator>
                    <MediaTime>1210 1300</MediaTime>
                40
                </VideoSegmentLocator>
            </HighlightLevel>
            <ImageLocator>
                <MediaTime>1270</MediaTime>
            45
            </ImageLocator>
        </HighlightSegment>
    </HighlightLevel>
    <Hierarchical summary>
    50
    55

```

[0053] The following is an example of a set of two highlights according to the above-described summary description schemes referring to particular events in a program, in particular "slam dunks" and "three-point shots" in a basketball game video. The first summary contains two video clips, each showing a slam-dunk; the second summary contains two video clips, each showing a three-point shot. By grouping the clips into summaries of events, a user may choose to view only the clips of slam-dunks; alternatively, the user may view all three-point shots. Note that in this case, there is no notion of hierarchy in the underlying real-world events.

```

10      <HierarchicalSummary name="mySummary" summaryType="highlight">
15          <HighlightLevel name="Slam dunks">
20              <HighlightSegment>
25                  <VideoSegmentLocator>
30                      <MediaTime>500 680</MediaTime>
35                  </VideoSegmentLocator>
40                  <ImageLocator>
45                      <MediaTime>590</MediaTime>
50                  </ImageLocator>
55              </HighlightSegment>
60          </HighlightLevel>
65          <HighlightLevel name="Three-point shots">
70              <HighlightSegment>
75                  <VideoSegmentLocator>
80                      <MediaTime>1200 1380</MediaTime>
85                  </VideoSegmentLocator>
90                  <ImageLocator>
95                      <MediaTime>1320</MediaTime>
100                  </ImageLocator>
105              </HighlightSegment>
110          </HighlightLevel>
115      </HierarchicalSummary>
120
125
130
135
140
145
150
155
160
165
170
175
180
185
190
195
200
205
210
215
220
225
230
235
240
245
250
255
260
265
270
275
280
285
290
295
300
305
310
315
320
325
330
335
340
345
350
355
360
365
370
375
380
385
390
395
400
405
410
415
420
425
430
435
440
445
450
455
460
465
470
475
480
485
490
495
500
505
510
515
520
525
530
535
540
545
550
555
560
565
570
575
580
585
590
595
600
605
610
615
620
625
630
635
640
645
650
655
660
665
670
675
680
685
690
695
700
705
710
715
720
725
730
735
740
745
750
755
760
765
770
775
780
785
790
795
800
805
810
815
820
825
830
835
840
845
850
855
860
865
870
875
880
885
890
895
900
905
910
915
920
925
930
935
940
945
950
955
960
965
970
975
980
985
990
995
1000
1005
1010
1015
1020
1025
1030
1035
1040
1045
1050
1055
1060
1065
1070
1075
1080
1085
1090
1095
1100
1105
1110
1115
1120
1125
1130
1135
1140
1145
1150
1155
1160
1165
1170
1175
1180
1185
1190
1195
1200
1205
1210
1215
1220
1225
1230
1235
1240
1245
1250
1255
1260
1265
1270
1275
1280
1285
1290
1295
1300
1305
1310
1315
1320
1325
1330
1335
1340
1345
1350
1355
1360
1365
1370
1375
1380
1385
1390
1395
1400
1405
1410
1415
1420
1425
1430
1435
1440
1445
1450
1455
1460
1465
1470
1475
1480
1485
1490
1495
1500
1505
1510
1515
1520
1525
1530
1535
1540
1545
1550
1555
1560
1565
1570
1575
1580
1585
1590
1595
1600
1605
1610
1615
1620
1625
1630
1635
1640
1645
1650
1655
1660
1665
1670
1675
1680
1685
1690
1695
1700
1705
1710
1715
1720
1725
1730
1735
1740
1745
1750
1755
1760
1765
1770
1775
1780
1785
1790
1795
1800
1805
1810
1815
1820
1825
1830
1835
1840
1845
1850
1855
1860
1865
1870
1875
1880
1885
1890
1895
1900
1905
1910
1915
1920
1925
1930
1935
1940
1945
1950
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
2015
2020
2025
2030
2035
2040
2045
2050
2055
2060
2065
2070
2075
2080
2085
2090
2095
2100
2105
2110
2115
2120
2125
2130
2135
2140
2145
2150
2155
2160
2165
2170
2175
2180
2185
2190
2195
2200
2205
2210
2215
2220
2225
2230
2235
2240
2245
2250
2255
2260
2265
2270
2275
2280
2285
2290
2295
2300
2305
2310
2315
2320
2325
2330
2335
2340
2345
2350
2355
2360
2365
2370
2375
2380
2385
2390
2395
2400
2405
2410
2415
2420
2425
2430
2435
2440
2445
2450
2455
2460
2465
2470
2475
2480
2485
2490
2495
2500
2505
2510
2515
2520
2525
2530
2535
2540
2545
2550
2555
2560
2565
2570
2575
2580
2585
2590
2595
2600
2605
2610
2615
2620
2625
2630
2635
2640
2645
2650
2655
2660
2665
2670
2675
2680
2685
2690
2695
2700
2705
2710
2715
2720
2725
2730
2735
2740
2745
2750
2755
2760
2765
2770
2775
2780
2785
2790
2795
2800
2805
2810
2815
2820
2825
2830
2835
2840
2845
2850
2855
2860
2865
2870
2875
2880
2885
2890
2895
2900
2905
2910
2915
2920
2925
2930
2935
2940
2945
2950
2955
2960
2965
2970
2975
2980
2985
2990
2995
3000
3005
3010
3015
3020
3025
3030
3035
3040
3045
3050
3055
3060
3065
3070
3075
3080
3085
3090
3095
3100
3105
3110
3115
3120
3125
3130
3135
3140
3145
3150
3155
3160
3165
3170
3175
3180
3185
3190
3195
3200
3205
3210
3215
3220
3225
3230
3235
3240
3245
3250
3255
3260
3265
3270
3275
3280
3285
3290
3295
3300
3305
3310
3315
3320
3325
3330
3335
3340
3345
3350
3355
3360
3365
3370
3375
3380
3385
3390
3395
3400
3405
3410
3415
3420
3425
3430
3435
3440
3445
3450
3455
3460
3465
3470
3475
3480
3485
3490
3495
3500
3505
3510
3515
3520
3525
3530
3535
3540
3545
3550
3555
3560
3565
3570
3575
3580
3585
3590
3595
3600
3605
3610
3615
3620
3625
3630
3635
3640
3645
3650
3655
3660
3665
3670
3675
3680
3685
3690
3695
3700
3705
3710
3715
3720
3725
3730
3735
3740
3745
3750
3755
3760
3765
3770
3775
3780
3785
3790
3795
3800
3805
3810
3815
3820
3825
3830
3835
3840
3845
3850
3855
3860
3865
3870
3875
3880
3885
3890
3895
3900
3905
3910
3915
3920
3925
3930
3935
3940
3945
3950
3955
3960
3965
3970
3975
3980
3985
3990
3995
4000
4005
4010
4015
4020
4025
4030
4035
4040
4045
4050
4055
4060
4065
4070
4075
4080
4085
4090
4095
4100
4105
4110
4115
4120
4125
4130
4135
4140
4145
4150
4155
4160
4165
4170
4175
4180
4185
4190
4195
4200
4205
4210
4215
4220
4225
4230
4235
4240
4245
4250
4255
4260
4265
4270
4275
4280
4285
4290
4295
4300
4305
4310
4315
4320
4325
4330
4335
4340
4345
4350
4355
4360
4365
4370
4375
4380
4385
4390
4395
4400
4405
4410
4415
4420
4425
4430
4435
4440
4445
4450
4455
4460
4465
4470
4475
4480
4485
4490
4495
4500
4505
4510
4515
4520
4525
4530
4535
4540
4545
4550
4555
4560
4565
4570
4575
4580
4585
4590
4595
4600
4605
4610
4615
4620
4625
4630
4635
4640
4645
4650
4655
4660
4665
4670
4675
4680
4685
4690
4695
4700
4705
4710
4715
4720
4725
4730
4735
4740
4745
4750
4755
4760
4765
4770
4775
4780
4785
4790
4795
4800
4805
4810
4815
4820
4825
4830
4835
4840
4845
4850
4855
4860
4865
4870
4875
4880
4885
4890
4895
4900
4905
4910
4915
4920
4925
4930
4935
4940
4945
4950
4955
4960
4965
4970
4975
4980
4985
4990
4995
5000
5005
5010
5015
5020
5025
5030
5035
5040
5045
5050
5055
5060
5065
5070
5075
5080
5085
5090
5095
5100
5105
5110
5115
5120
5125
5130
5135
5140
5145
5150
5155
5160
5165
5170
5175
5180
5185
5190
5195
5200
5205
5210
5215
5220
5225
5230
5235
5240
5245
5250
5255
5260
5265
5270
5275
5280
5285
5290
5295
5300
5305
5310
5315
5320
5325
5330
5335
5340
5345
5350
5355
5360
5365
5370
5375
5380
5385
5390
5395
5400
5405
5410
5415
5420
5425
5430
5435
5440
5445
5450
5455
5460
5465
5470
5475
5480
5485
5490
5495
5500
5505
5510
5515
5520
5525
5530
5535
5540
5545
5550
5555
5560
5565
5570
5575
5580
5585
5590
5595
5600
5605
5610
5615
5620
5625
5630
5635
5640
5645
5650
5655
5660
5665
5670
5675
5680
5685
5690
5695
5700
5705
5710
5715
5720
5725
5730
5735
5740
5745
5750
5755
5760
5765
5770
5775
5780
5785
5790
5795
5800
5805
5810
5815
5820
5825
5830
5835
5840
5845
5850
5855
5860
5865
5870
5875
5880
5885
5890
5895
5900
5905
5910
5915
5920
5925
5930
5935
5940
5945
5950
5955
5960
5965
5970
5975
5980
5985
5990
5995
6000
6005
6010
6015
6020
6025
6030
6035
6040
6045
6050
6055
6060
6065
6070
6075
6080
6085
6090
6095
6100
6105
6110
6115
6120
6125
6130
6135
6140
6145
6150
6155
6160
6165
6170
6175
6180
6185
6190
6195
6200
6205
6210
6215
6220
6225
6230
6235
6240
6245
6250
6255
6260
6265
6270
6275
6280
6285
6290
6295
6300
6305
6310
6315
6320
6325
6330
6335
6340
6345
6350
6355
6360
6365
6370
6375
6380
6385
6390
6395
6400
6405
6410
6415
6420
6425
6430
6435
6440
6445
6450
6455
6460
6465
6470
6475
6480
6485
6490
6495
6500
6505
6510
6515
6520
6525
6530
6535
6540
6545
6550
6555
6560
6565
6570
6575
6580
6585
6590
6595
6600
6605
6610
6615
6620
6625
6630
6635
6640
6645
6650
6655
6660
6665
6670
6675
6680
6685
6690
6695
6700
6705
6710
6715
6720
6725
6730
6735
6740
6745
6750
6755
6760
6765
6770
6775
6780
6785
6790
6795
6800
6805
6810
6815
6820
6825
6830
6835
6840
6845
6850
6855
6860
6865
6870
6875
6880
6885
6890
6895
6900
6905
6910
6915
6920
6925
6930
6935
6940
6945
6950
6955
6960
6965
6970
6975
6980
6985
6990
6995
7000
7005
7010
7015
7020
7025
7030
7035
7040
7045
7050
7055
7060
7065
7070
7075
7080
7085
7090
7095
7100
7105
7110
7115
7120
7125
7130
7135
7140
7145
7150
7155
7160
7165
7170
7175
7180
7185
7190
7195
7200
7205
7210
7215
7220
7225
7230
7235
7240
7245
7250
7255
7260
7265
7270
7275
7280
7285
7290
7295
7300
7305
7310
7315
7320
7325
7330
7335
7340
7345
7350
7355
7360
7365
7370
7375
7380
7385
7390
7395
7400
7405
7410
7415
7420
7425
7430
7435
7440
7445
7450
7455
7460
7465
7470
7475
7480
7485
7490
7495
7500
7505
7510
7515
7520
7525
7530
7535
7540
7545
7550
7555
7560
7565
7570
7575
7580
7585
7590
7595
7600
7605
7610
7615
7620
7625
7630
7635
7640
7645
7650
7655
7660
7665
7670
7675
7680
7685
7690
7695
7700
7705
7710
7715
7720
7725
7730
7735
7740
7745
7750
7755
7760
7765
7770
7775
7780
7785
7790
7795
7800
7805
7810
7815
7820
7825
7830
7835
7840
7845
7850
7855
7860
7865
7870
7875
7880
7885
7890
7895
7900
7905
7910
7915
7920
7925
7930
7935
7940
7945
7950
7955
7960
7965
7970
7975
7980
7985
7990
7995
8000
8005
8010
8015
8020
8025
8030
8035
8040
8045
8050
8055
8060
8065
8070
8075
8080
8085
8090
8095
8100
8105
8110
8115
8120
8125
8130
8135
8140
8145
8150
8155
8160
8165
8170
8175
8180
8185
8190
8195
8200
8205
8210
8215
8220
8225
8230
8235
8240
8245
8250
8255
8260
8265
8270
8275
8280
8285
8290
8295
8300
8305
8310
8315
8320
8325
8330
8335
8340
8345
8350
8355
8360
8365
8370
8375
8380
8385
8390
8395
8400
8405
8410
8415
8420
8425
8430
8435
8440
8445
8450
8455
8460
8465
8470
8475
8480
8485
8490
8495
8500
8505
8510
8515
8520
8525
8530
8535
8540
8545
8550
8555
8560
8565
8570
8575
8580
8585
8590
8595
8600
8605
8610
8615
8620
8625
8630
8635
8640
8645
8650
8655
8660
8665
8670
8675
8680
8685
8690
8695
8700
8705
8710
8715
8720
8725
8730
8735
8740
8745
8750
8755
8760
8765
8770
8775
8780
8785
8790
8795
8800
8805
8810
8815
8820
8825
8830
8835
8840
8845
8850
8855
8860
8865
8870
8875
8880
8885
8890
8895
8900
8905
8910
8915
8920
8925
8930
8935
8940
8945
8950
8955
8960
8965
8970
8975
8980
8985
8990
8995
9000
9005
9010
9015
9020
9025
9030
9035
9040
9045
9050
9055
9060
9065
9070
9075
9080
9085
9090
9095
9100
9105
9110
9115
9120
9125
9130
9135
9140
9145
9150
9155
9160
9165
9170
9175
9180
9185
9190
9195
9200
9205
9210
9215
9220
9225
9230
9235
9240
9245
9250
9255
9260
9265
9270
9275
9280
9285
9290
9295
9300
9305
9310
9315
9320
9325
9330
9335
9340
9345
9350
9355
9360
9365
9370
9375
9380
9385
9390
9395
9400
9405
9410
9415
9420
9425
9430
9435
9440
9445
9450
9455
9460
9465
9470
9475
9480
9485
9490
9495
9500
9505
9510
9515
9520
9525
9530
9535
9540
9545
9550
9555
9560
9565
9570
9575
9580
9585
9590
9595
9600
9605
9610
9615
9620
9625
9630
9635
9640
9645
9650
9655
9660
9665
9670
9675
9680
9685
9690
9695
9700
9705
9710
9715
9720
9725
9730
9735
9740
9745
9750
9755
9760
9765
9770
9775
9780
9785
9790
9795
9800
9805
9810
9815
9820
9825
9830
9835
9840
9845
9850
9855
9860
9865
9870
9875
9880
9885
9890
9895
9900
9905
9910
9915
9920
9925
9930
9935
9940
9945
9950
9955
9960
9965
9970
9975
9980
9985
9990
9995
9999

```

```

<MediaTime>2500 2680</MediaTime>
5 </VideoSegmentLocator>
<ImageLocator>
<MediaTime>2590</MediaTime>
</ImageLocator>
</HighlightSegment>
10 <HighlightSegment>
<VideoSegmentLocator>
<MediaTime>3200 3380</MediaTime>
</VideoSegmentLocator>
<ImageLocator>
15 <MediaTime>3320</MediaTime>
</ImageLocator>
</HighlightSegment>
</HighlightLevel>
20 </Hierarchical summary>

```

25 [0054] Having seen several examples of the hierarchical summary description scheme and its related components, it is helpful to look at the sequential summary description scheme 56 shown in Figure 4. The sequential summary description scheme is used to specify summaries of an audio-visual item consisting of an arbitrary but predetermined sequence of still images or video frames, which can be visualized sequentially in time. The playback speed of video frames can be controlled to enable smart fast-forwarding.

30 [0055] A sequential summary refines a summary and contains a single audio-visual summary. It contains either a sequence of references to still images, or a sequence of video frames. A sequential summary may contain a sequence of references to audio-clips. Audio-clips may be played back in synchronization with the video frames.

[0056] The following is an example of a simple sequential summary according to the above-described summary description schemes, representing an animated slide-show. It refers to a number of images, which may be shown in 35 sequential fashion, or under control of the user.

```

(SequentialSummary name="mySummary" summaryType="sequential")
40   (<ImageLocator><MediaUrl>file://images/photo1.jpg</MediaUrl></ImageLocator>
   (<ImageLocator><MediaUrl>file://images/photo2.jpg</MediaUrl></ImageLocator>
   (<ImageLocator><MediaUrl>file://images/photo3.jpg</MediaUrl></ImageLocator>
   (<ImageLocator><MediaUrl>file://images/photo4.jpg</MediaUrl></ImageLocator>
   (<ImageLocator><MediaUrl>file://images/photo5.jpg</MediaUrl></ImageLocator>
   (<ImageLocator><MediaUrl>file://images/photo6.jpg</MediaUrl></ImageLocator>
45   /</SequentialSummary>

```

50 [0057] All of these summaries can be presented as alternatives to the user. The user selects the type of summary desired, based upon the type of media representation desired and the level of specificity. As discussed in detail above, the representation can be one of several different types, with multiple levels and can be either dependent or independent.

[0058] The presentation of these summaries as well as the transfer and communications between the various entities involved in this presentation is shown in Figure 6. These are achieved by using common description schemes such as the above-mentioned title descriptions and summary descriptions. The content creator/provider 62 provides the 55 audiovisual programs and other data services (metadata) associated with those programs to the service provider 64. The data services may include such things as directories of key clips, or other types of indexes of the audiovisual program, for example, such as indexes to segments containing touchdowns and field goals in a football game. The service provider 64 may originally prepare summary description and text information in respect to an audiovisual program pos-

sessed by the content provider 62. The service provider 64 and the content provider 62 negotiate some type of fee arrangement for this transfer. The content provider 62 may also be the service provider 64. The user can select an audiovisual program based on the metadata such as title and summary descriptions provided by the service provider 64. The metadata given to the audiovisual program has the same structure as data used for the title and summary descriptions.

5 Hence, an audiovisual program may have metadata provided from plural service providers. The user at the system side can adaptively obtain a desired content by his/her preference for the summary and as the result of comparison of means, expenses and time duration necessary for obtaining the program by utilizing a variety of summaries for the same audiovisual programs.

[0059] The service provider 64 then sends the summary description using the above-described summary description scheme about the various audiovisual programs to the user 60. The service provider also tracks the resources at the user's end and the user preference and history. Tracking such information is desirable for the service provider to offer the user descriptions for summaries that are desirable to the user and usable by user's platform. The user can then receive summary descriptions according to the users preference that operate on the content provided by the content provider. There is again some fee arrangement between the service provider and the user.

10 [0060] The content provider may also track the user preferences and usage history to directly deliver the summarized programs to the user 60. In this case, the summary descriptions reside at the content provider, and the content provider selects, according to user data, generates and directly delivers the appropriate program summary to the user.

15 [0061] In this particular example, the user transacts separately with the content provider and the service provider for content and summary services, respectively. However, all the functionality provided within box 70 could be provided by either the service provider or the content provider. Some content providers may decide to offer these services, as well as some service providers deciding to offer content. It is also possible that the user has arrangements with other providers. For one service or type of content the two providers could be separated, for other services or types of content, the two could be combined together. In this case, where the service and content providers are the same, the user preferences and usage history would be sent only once.

20 [0062] The user may also interact with other users to exchange information by using the above-described description scheme. For example, the user may have the capability to produce customized audiovisual program summaries (e.g., "My Favorite Moments") at the user end. The user could then pass these customized summary descriptions to other users 66 to share experiences or to make reviews and recommendations about a particular program. Other users could then receive summary descriptions that operate on these programs provided by a content provider. Alternatively,

25 one user transfers only a description data of a customized summary for an audiovisual program to the other user that can then directly refer to and view the audiovisual program specified by the customized summary.

30 [0063] In this manner, the descriptive framework is used to provide summary descriptions. These summary descriptions can then be used to present alternative summaries of audiovisual content to the user. The content and the summary descriptions are provided according to an arrangement of transfers and transactions.

35 [0064] Thus, although there has been described to this point a particular embodiment for a method and structure for as description framework for audiovisual presentation systems, it is not intended that such specific references be considered as limitations upon the scope of this invention except in-so-far as set forth in the following claims.

Claims

40 1. A system operable to provide a description framework about programs presented by an audiovisual system, the framework including a descriptive structure operable to identify and locate each of the audiovisual representations of audiovisual material.

45 2. The system of claim 1, wherein the description framework includes at least one multimedia title description.

3. The system of claim 2, wherein the title description included in the description framework collectively describes text data of a title of an audiovisual content and multimedia data representing a content of at least one audiovisual program.

50 4. The system of claim 2, wherein the title description included in the description framework includes information for locating multimedia data and is capable of handling multimedia data being outside the audiovisual system.

55 5. The system of claim 1, wherein the description framework includes at least one summary description.

6. The system of claim 5, wherein the summary description included in the description framework comprises a combination of at least one piece of partial data obtainable by extracting a part of multimedia data composing the audiovisual content and is described using information for locating the multimedia data and the partial data extracted

therefrom.

- 5 7. The system of claim 5, wherein the summary description included in the description framework includes multimedia data and information for locating to extract a partial multimedia data from the multimedia data and capable of handling multimedia data being outside the audiovisual system.
8. The system of claim 1, wherein the description framework includes one of either a multimedia title description or a summary description and at least one other component.
- 10 9. A method of presenting summaries of audiovisual content to a user, the method comprising the steps of:
 - a) presenting a multi-view menu of the available types of summaries to the user, wherein the multi-view can provide hierarchical and non-hierarchical summaries;
 - b) receiving a user selection of a summary type; and
 - c) providing summaries of the selected type to the user.
- 15 10. A method of providing summary description services of audiovisual content to a user (60), comprising the steps of:
 - a) receiving information from a user (60), wherein said information includes specifications of platform resources at the user end and user preferences;
 - b) tracking usage history of the user (60);
 - c) transmitting audiovisual material to the user (60); and
 - d) providing summary descriptions operable to be applied to the audiovisual material to the user (60).
- 20 11. The method of claim 10 wherein said receiving and providing steps are performed by a service provider (64).
12. The method of claim 11 wherein said transmitting audiovisual content step is performed by a content provider (62).
- 25 13. The method of claim 10 wherein the steps are performed by a combination content and service provider (62, 64).
- 30 14. The method of claim 10 wherein the summary descriptions provided to the user (60) are of a format allowing the user (60) to exchange summary descriptions with other users (66).
- 35 15. The method of claim 14 wherein the format allowing the user (60) to exchange summary descriptions also allows the user (60) to customize summary descriptions.

40

45

50

55

FIG.1

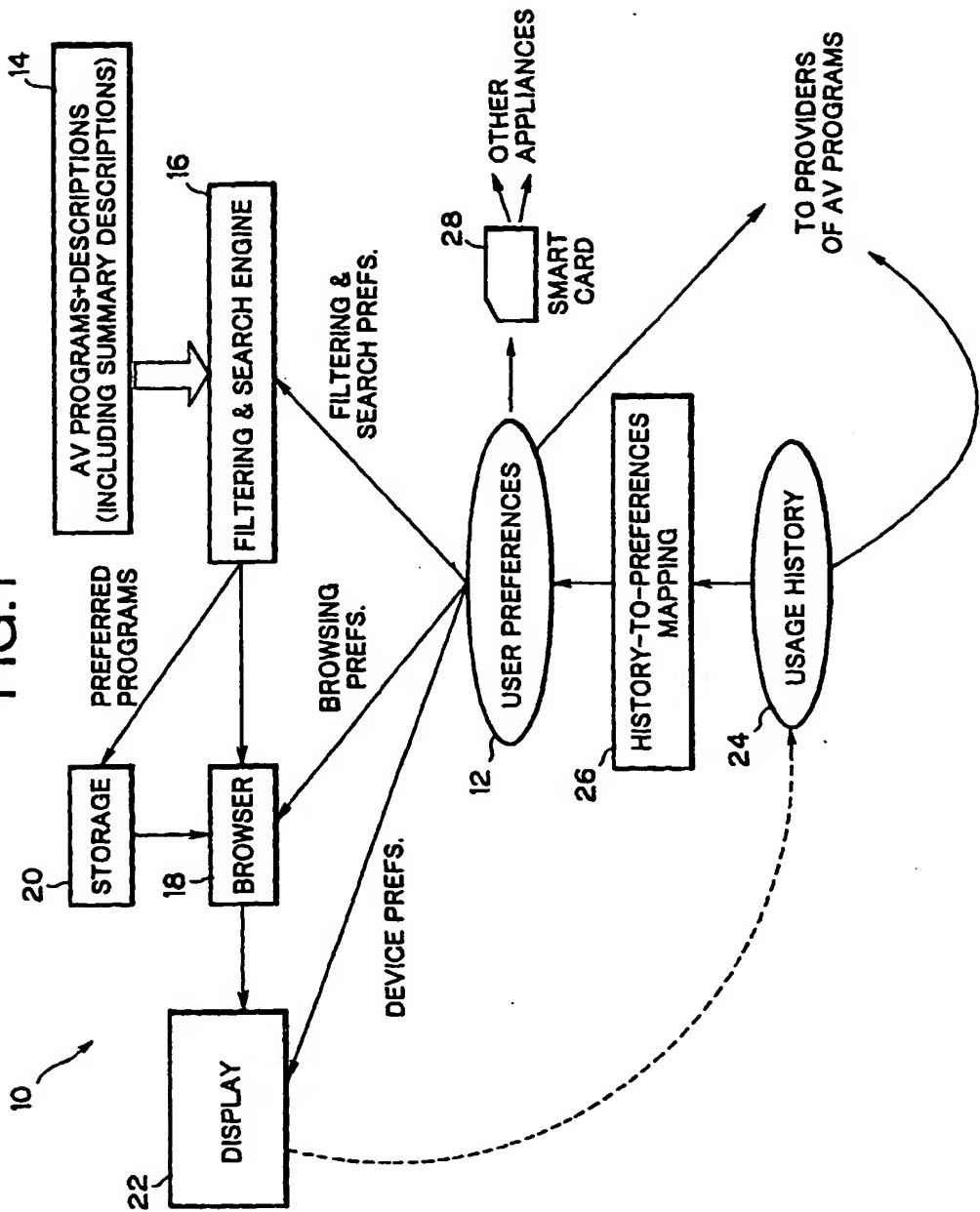


FIG.2

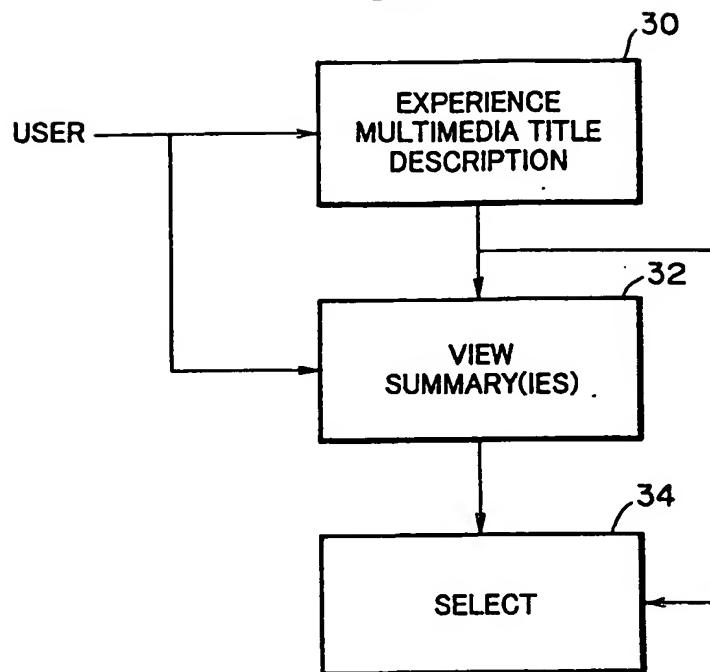


FIG.3

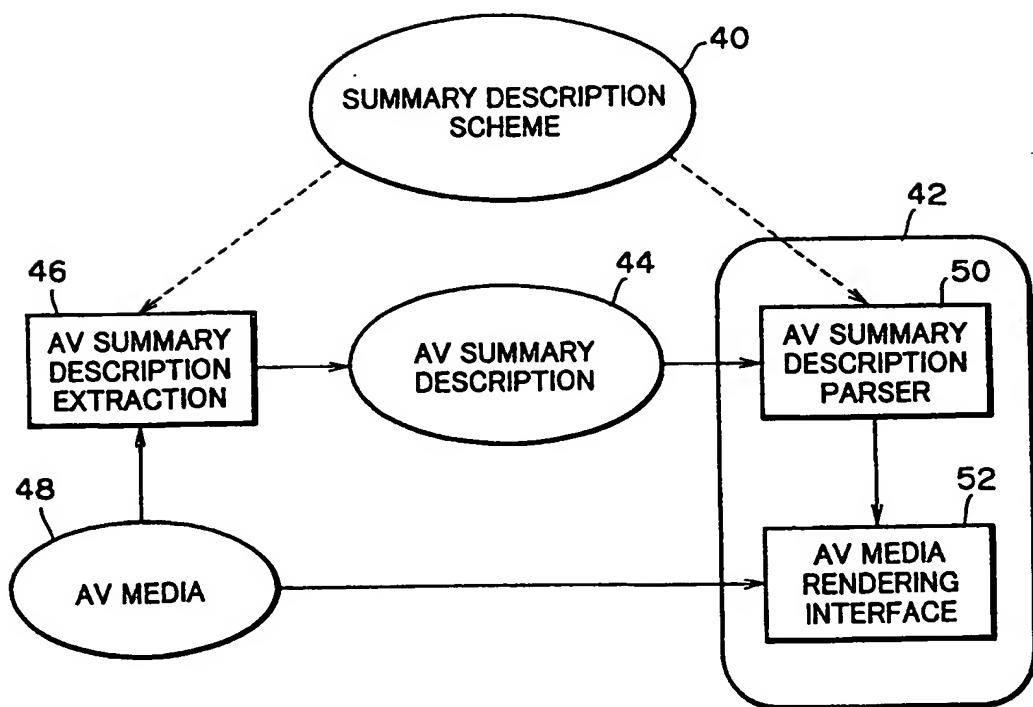


FIG.4

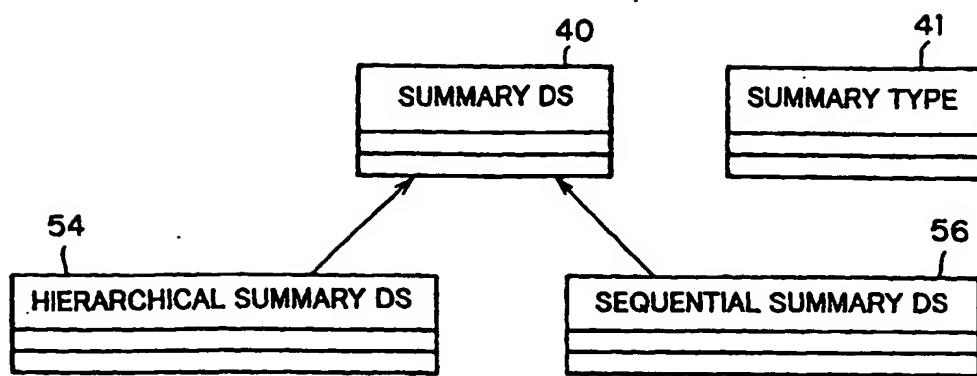


FIG.5

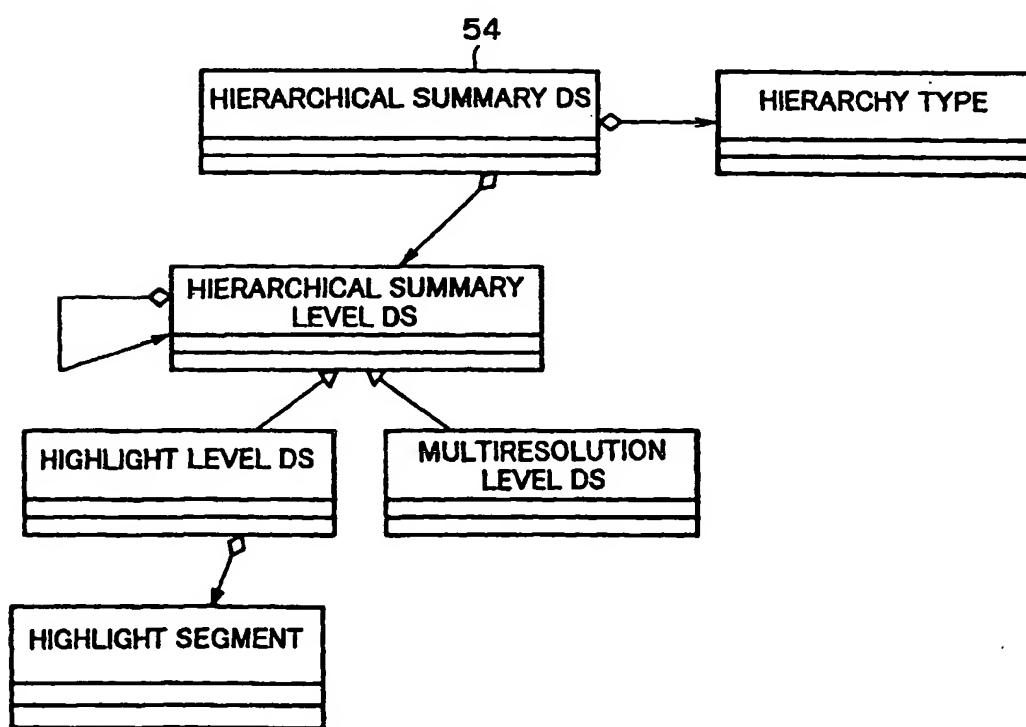


FIG.6

